

다중 PCA모듈을 이용한 얼굴포즈 판별

고재필*, 김선욱, 변혜란
연세대학교 컴퓨터 과학과

Multiple PCA Module Based Face Pose Estimation

Jaephil Ko, SunWook Kim, Hyeran Byun
Dept. of Computer Science, Yonsei University

요 약

본 논문에서는 얼굴인식에 주로 사용되는 PCA를 얼굴포즈판별에 적용해 보았다. 얼굴포즈판별은 개개인의 얼굴특징을 강조해야 하는 얼굴인식과는 달리 일반적인 얼굴특징을 이용하기 때문에 PCA에 적합한 응용분야이다. 그러나, 다양한 얼굴포즈에 대한 영상을 하나의 표본집합으로 사용하면, 표본집합의 분산이 크기때문에 포즈별로 표본집합을 달리하여 PCA모듈을 구성하는 것이 타당하다. 표본수집의 어려움은 3차원 한국인 표준모형을 이용해 극복하고, 이를 통하여 다양한 조명방향 및 얼굴포즈에 대한 표본을 수집하였다. 5방향의 얼굴포즈에 대한 판별 실험을 통하여 모듈화된 PCA의 분류기로서의 가능성을 살펴보고, 조명에 따른 오류를 완화하고자 비 선형적 패턴을 나타내는 각 PCA모듈의 결과를 신경망에 적용하여 보았다.

1. 서 론

최근의 HCI가 사용자에게 사용하기 편리한 환경을 제공하기 위한 연구를 진행중이며, 이의 일환으로 얼굴인식은 생체인식 방법 중 가장 사용자에게 친숙하다라는 장점으로 많은 연구가 수행되어 왔다. 이 중 PCA를 이용하는 방법은 고차원으로 표현되는 얼굴특징을 효과적으로 저차원으로 줄여 주기 때문에 매칭에 매우 유리하다[1].

PCA를 이용하는 방법에는 크게 두 가지가 있다. 그림 1-(a)에서와 같이 PC를 이용하는 방법과 복원영상을 이용하는 방법이다. 얼굴인식은 인식할 대상이 되는 모든 얼굴을 하나의 표본집합으로 하여 PCA를 적용하고 데이터베이스에 PC를 저장하여 사용한다. 얼굴영역 추출, 포즈추정등에는 복원영상을 사용한다[2].

표본영상간의 분산이 큰 얼굴포즈인 경우에 고유벡터를 적은 수만 사용하여 영상을 복원하면 원영상과 비교시 에러가 커진다. 그렇다고, 사용할 고유벡터의 수를 높이면 개개인의 얼굴특성을 반영하게 되어 일반적인 얼굴의 특징을 사용하게 되는 포즈추정에는 적합하지 않다. 따라서, 포즈별로 PCA모듈을 구성하는 것이 좋다.

본 논문에서는 왼쪽에서 오른쪽으로 5가지 얼굴방향을 설정하여 표본을 수집하고, 이들에 대하여 PCA모듈을 구성한다. 새로운 얼굴영상과 5개의 PCA모듈을 통해 복원한 영상을 비교하여 이 중 오차가 가장 작은 복원영상을 출력한 PCA모듈의 얼굴포즈를 입력영상의 포

즈로 결정한다.

다양한 표본수집 특히, 빛의 방향에 따른 표본 획득의 어려움은 3차원 모델을 통한 합성방안을 따랐다[3]. 본 논문에서는 3차원 한국인 표준모형[4]을 통해 하나의 얼굴영상에 대해 광원을 9군데로 변경시키면서 표본을 수집하였다.

빛에 의한 오판을 줄이고자 포즈별 PCA모듈을 통해 나온 결과를 신경망에 학습하여 포즈분류 실험을 수행하였다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 PCA모듈을 설명하고, 3장에서는 다양한 실험을 통해 분류기로서의 가능성을 살펴보고, 마지막 4장에서는 결론을 짓는다.

2. PCA모듈

2.1. PCA(Principle Component Analysis)[5]

PCA는 데이터가 가지고 있는 원특징을 가지고 있으면서 효율적으로 차원을 줄일 수 있다. 차원이 $n \times n$ 인 랭덤벡터 X 를 평균이 0이라고 가정하면 즉, $E[X] = 0$ 일 때 다음과 같이 X 에 대한 $n \times n$ 차원의 공분산 행렬 R 은 벡터 X 에 대한 외적의 기대값으로 정의된다.

$$R = E[XX^T] \quad (1)$$

공분산 행렬 R 의 고유값들을 $\lambda_1, \lambda_2, \dots, \lambda_n$, 대응되는 고유벡터들을

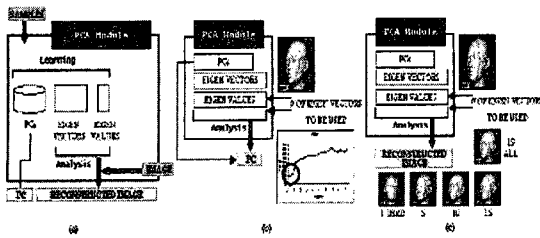


그림 1 PCA 모듈

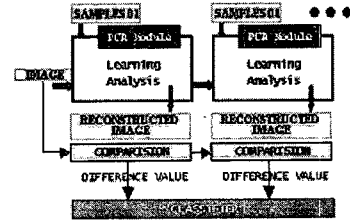


그림 2 다중 PCA모듈

a_1, a_2, \dots, a_n 라고 하면 다음과 같이 쓸 수 있다.

$$Rq_j = \lambda_j q_j \quad j=1, 2, \dots, n \quad (2)$$

여기서 고유값들이 다음과 같은 순서로 정렬되어 있다고 하자.

$$\lambda_1 > \lambda_2 > \dots > \lambda_j > \dots > \lambda_n \quad (3)$$

그리고 대응되는 고유벡터를 다음과 같이 표현하자.

$$Q = [a_1, a_2, \dots, a_n] \quad (4)$$

그러면, PCA는 X 의 어떤 실제벡터를 x 라하고 변환행렬로 Q 를 사용하여 특징벡터 y 를 얻는 다음 식으로 정의된다.

$$y = [y_1, y_2, \dots, y_n]^T = [x^T a_1, x^T a_2, \dots, x^T a_n]^T = Q^T x \quad (5)$$

$$y_j = a_j^T x = x^T a_j, \quad j=1, 2, \dots, n \quad (6)$$

x 를 단위벡터 a_j 로 표현되는 주요방향으로 투영하여 얻은 결과를 y_j 라고 할 때 이를 주성분이라고 한다. 데이터 벡터 x 는 특징벡터 y 로부터 다음과 같이 복원 가능하다.

$$x = \sum_{j=1}^n y_j a_j = Qy \quad (7)$$

여기서 값이 큰 l 개의 λ 만을 사용하여 복원한 \hat{x} 와 x 의 오차 e 는 다음 수식과 같고, q 벡터들은 서로 수직이기 때문에 에러값은 l 을 증가시키에 따라 급격하게 떨어진다.

$$e = x - \hat{x} = \sum_{j=l+1}^n \lambda_j \quad (8)$$

2.2. PCA 모듈 기능 확인

모듈기능을 확인하기 위하여 19장의 얼굴포즈영상을 표본영상으로 사용하였다. 그림 1-(b)는 얼굴인식에서 사용하는 방법으로 표본에서 사용하지 않은 좌측을 향하고 있는 영상을 입력으로 하였을 때 표본에서 제일 좌측 PC와 유사도가 가장 큼을 확인 할 수 있다. 그림 1-(c)는 복원영상을 보여주는 그림으로 사용하는 고유벡터 수가 증가할 수록 입력영상과 유사해 진다.

2.3. 얼굴포즈를 판별하기 위한 다중 PCA 모듈

포즈별 PCA모듈을 그림 2와 같이 구성한다. 새로운 입력영상과 이를 각 모듈에 적용하여 얻은 복원영상과의 픽셀간 거리를 얻는다. 이때 분류기에서는 가장 작은 값을 출력한 모듈을 승자로 선택 할 수도 있으며, 또는 출력값들을 입력으로 하는 신경망을 사용하여 승자를 결정할 수도 있다.

3. 실험

다중 PCA모듈의 성능을 실험하기 위하여 3차원 모델로부터 그림3과 같은 얼굴표본을 수집하였다. 데이터집합은 2가지로 ㉠시선의 상중화, 5포즈, 포즈당 5장면을 선정한 총 75장의 영상과, ㉡시선의 상중화, 5포즈, 포즈당 1장면, 장면당 광원 9군데로 총 135장의 영상을 수집하였다. 광원은 영상을 가로세로 3등분하여 교차하는 지점으로 설정하였다. 표본영상은 ㉠에서 시선이 가운데인 표본을 사용하여 포즈별 5장을 사용하여 PCA모듈을 구성하였다.



그림 3 3차원 모델로부터 생성한 얼굴표본

3.1. 표본영상과 입력영상의 PC값 및 복원영상 비교

각 모듈의 PC값과 학습에 사용하지 않은 영상의 PC값 비교결과는 그림 4에 제시하였다. 그림 4-(e)에서 보듯이 입력영상 R2는 표본영상 R2보다 오히려 표본영상 R1과 더 유사하다. 그러나, 그림 5을 보면, 그림 4-(e)에서 보인 문제는 복원영상을 이용한 경우 나타나지 않는다. 이로서 다중 PCA모듈을 이용하는 경우에 원 영상과 복원영상간의 차이를 이용하는 것이 단순히 PC를 비교하는 것 보다 성능이 낫다고 볼 수 있다.

3.2. 분류기 성능

포즈분류성능은 3.1절에서 보았듯이 입력영상과 복원영상간의 픽셀간 차이를 사용하였다. 가장 작은 값을 출력한 모듈을 승자로 결정하는 방법과, 이들 출력값의 비 선형적 패턴을 신경망에 입력하여 승자를 결정하는 방법에 대하여 데이터집합 ㉠에 대하여 수행하였다.

먼저, 가장 작은 값을 출력하는 모듈을 승자로 결정하는 방법에 대한 실험결과는 데이터집합㉠, 75장의 영상 모두에 대하여 성공하였고, 광원의 방향이 다른 데이터집합㉡, 135장의 영상에 대해서는 5장의 오류가 발생하여 96.3%의 성공률을 보였다. 데이터집합㉡는 PCA모듈 생성시 사용한 표본집합과는 다른 광원을 가졌음에도 불구하고 비교적 높은 성공률을 보여주었다.

각 모듈이 출력하는 차이값의 최소값으로만 승자를 결정하는 경우의 오류를 줄여 보고자, 5개의 출력값들이 나타내는 비선형적 패턴을 신경망에 학습하여 성능을 실험하였다. 신경망은 3계층으로 입력, 히

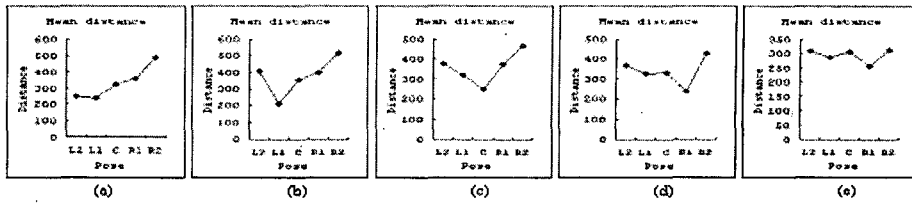


그림 4 PC값 비교

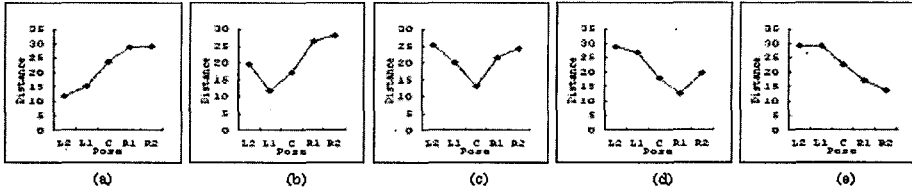


그림 5 복원영상 비교

든, 출력노드 각각 5, 7, 5개로 구성하였으며, 학습알고리즘은 교사학습알고리즘의 대표인 백프로퍼게이션을 사용하였다. 데이터집합④⑥를 합한 총 210장의 영상에 대하여 각 모델이 출력하는 5개의 차이값을 입력으로 하였으며, 10-fold-cross-validation방법[6]으로 실험하여, 최고 98.57%의 성공율을 보여 주었다. 그러나, 신경망을 이용한 경우, 학습에 사용한 표본영상에 크게 의존하였다.

4. 결 론

표본집합의 분산이 비교적 큰 경우인 얼굴포즈를 판별하기 위해 PCA를 이용하는 경우, 포즈별로 PCA모듈을 구성하는 것이 타당하다. 이때는 복원영상을 사용하는 것이 효과적임이 실험을 통하여 확인하였다.

얼굴위치나, 얼굴포즈와 같이 개개의 얼굴특징이 아닌 일반적인 얼굴형태를 특징으로 하는 경우, PCA는 큰 잇점을 가진다. 왜냐하면, 표본집합의 주요한 특징을 뽑아내기 때문이다. 이러한 특징은 광원의 위치를 달리하여 얼굴포즈를 판별하였을 때 96.3%의 높은 성공률을 보인 것으로도 반증된다. 왜냐하면, PCA가 광원의 위치에 따른 얼굴명암도 변화를 예러로 보고 전체적인 얼굴포즈의 특징을 잃지 않는다는 것을 말하기 때문이다.

PCA의 또 다른 장점은 신경망과 같이 네거티브한 표본을 사용하지 않아도 된다는 점으로, 얼굴포즈와 같이 몇 가지로 제한된 분류를 구분해 내는 응용에 다중 PCA모듈은 효과적으로 사용될 수 있을 것이다.

본 논문에서 표본영상으로 사용한 3차원 한국인 모델은 한국인의 표준적인 얼굴을 나타내므로, 3차원 모델에서 생성한 표본으로 구성된 PCA모듈로 실제 한국인의 얼굴포즈 판별에도 적용 가능하다.

5 참고 문헌

[1] Matthew A.Turk and Alex P. Pentland, "Face Recognition Using Eigenfaces", Proc. of CVPR,

pp.586-591, 1991

[2] Gong, S, McKenna, S, Collins, J.J. "An investigation into face pose distributions", Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, pp.265 -270, 1996

[3] Satio H, Watanabe A, Ozawa S "Face pose estimating system based on eigenspace analysis", Image Processing, pp.638-642, 1999

[4] 심연숙, 김선옥, 한재현, 변혜란, 정찬섭, "자연스러운 표정 합성을 위한 3차원 얼굴 모델링 및 합성 시스템", 한국인지과학회, 제11권 제2호, 2000.

[5] Haykin, Neural Networks, A Comprehensive Foundation , Prentice Hall, pp.392-442

[6] Christopher M. Bishop, "Neural Networks for Pattern Recognition", Clarendon Press. Oxford, 1995